High-content siRNA screening

Eberhard Krausz*

DOI: 10.1039/b616187c

Very recent developments in instrumentation and image analysis have made microscopy applicable to high-throughput screening (HTS). For 'High-Content Screening' modern automated microscopy systems provide a throughput of up to 100 000 (confocal) images, with amazingly high resolution, of cells fluorescently stained using multiple colours that are imaged simultaneously during the screen. Image analysis tools provide multi-parametric pattern extraction and quantification on-the-fly. Big pharmaceutical companies have presented image-based screens of more than 100 000 compounds, while academia has published data on large RNA interference screens for functional genomics.

Numerous whole-genome sequencing projects have been completed and published. Gene annotation is still in flux. Nevertheless, about 23 000 human genes have been reliably annotated. Additionally, gene expression array technologies and proteomics have added further data on molecules present in cells and tissues. The major challenge of the present and future is to unravel the detailed function of all these gene products and their interaction. One way to gain insight, is to design oligonucleotides that induce lack-of-function phenotypes by specifically inhibiting protein production.

What is ''High-content screening''?

The daily work of cell biologists is based on microscopy. Particularly, fluorescent microscopy has enabled multifaceted insights into the detail and complexity of cellular structures and their functions for well over two decades. As an essential prerequisite for a systematic phenotypical analysis of gene functions in cells at a genome-wide scale, the throughput of microscopy had to be improved through automation. Along with the introduction of the first automated fluorescent imaging systems in the late 90's the term 'High-Content Screening' (HCS) was coined. HCS is defined as multiplexed functional screening based on imaging multiple targets in the physiologic context of intact cells by extraction of multicolour fluorescence information.[1]

There are a number of advantages of high-content screening over other screening technologies. First of all, cell-based assays reflect high physiological relevance; particularly with regard to drug-screening, the response is not limited to a single target but rather to a whole cell containing thousands of targets. Putative cytotoxic (side-) effects are discovered early on. Secondly, single cell analysis reflects the cell populations' heterogeneity as well as their individual response to treatment. Simultaneous staining in 3 or 4 colours allows the extraction of various parameters from each cell quantitatively as well as qualitatively such as intensity, size, distance or distribution (spatial resolution). The parameters might be referenced to each other, for example the use of nuclei staining to normalize other signals against cell number, or particular parameters might verify or exclude each other. Generally, high-content analysis (HCA) is considered to produce low false-positive and false-negative results.

On the other hand, there are a number of challenges and some significant limitations. In image-based cellular assays higher variability is observed compared to biochemical assays or homogeneous cell-based assays such as reporter gene assays or cytoblots[2] whereby the variability across the cell layer is normalized

Head, HT-Technology Development Studio (TDS), Max Planck Institute of Molecular Cell Biology and Genetics (MPI-CBG), Pfotenhauerstrasse 108, D-01307 Dresden, Germany. E-mail: krausz@mpi-cbg.de; http://tds.mpi-cbg.de/webtds/; Fax: +49 351 210 - 1349; Tel: +49 351 210 - 2732

Eberhard Krausz

Eberhard Krausz is currently heading the 'High Throughput - Technology Development Studio' (TDS), the High-Content Screening Operations at the Max Planck Institute of Molecular Cell Biology and Genetics (MPI-CBG) in Dresden, Germany. Previously he held responsible positions for 6 years in the biotech industry at Cenix BioScience GmbH (Dresden, Germany) and Cyclacel Ltd. (Dundee, Scotland/UK) dealing with RNA interference for target identification and validation, and drug discovery targeting kinases, respectively. Previously he ran a gene therapy project in restenosis in a medical research centre in Munich. He earned his PhD at the LMU München/Munich studying gene regulation in eye lens development. He is a microbiologist by training.

by lysing 10 000 or more cells. However, in HCA variability is a more critical issue as typically only segments of the whole well are selected to analyse about 200–1000 cells, and so this is highly dependent on even distribution and minimal loss of cells particularly during washing processes.

"High-throughput" in HCS is relative: although instruments that acquire in the range of 100 000 images per day are already marketed, this is still not comparable to the throughput of classical HTS. Assays get more and more complex, consequently assay development times become prolonged. Further, standardisation of cell culture conditions is a major challenge. Major investments have to be made to gather a critical mass of instrumentation, image analysis tools and IT infrastructure. The data load per run of a screen may easily go beyond the one Terabyte border, and the processing of the hundreds of thousands of images applying complex image analysis software and algorithms requires an extraordinarily powerful IT infrastructure. In this article we will give an overview of the considerations that should be kept in mind while setting-up the infrastructure to implement and successfully run large-scale high-content siRNA screens for functional genomics.

RNA interference molecule libraries

RNA interference (RNAi) is a natural mechanism[3,4] to specifically silence individual genes. Researchers exploit this mechanism to identify and characterize the molecular functions of genes in the context of a living cell. By sequence-specific degradation of a desired mRNA population in a cell by short interfering RNA (siRNA) molecules the synthesis of the encoded protein is diminished. Sophisticated algorithms have been developed to predict functional siRNA molecules at high probability,[5–7] and siRNAs are commercially available against all predicted genes of the genomes of human, mouse and rat.[8] An increasing list of so-called "validated" siRNAs that have been tested for efficient mRNA degradation in cells by the manufacturers is being generated. Meanwhile larger collections of siRNA molecules are available to target whole gene families, target classes such as kinases or phosphatases, the so-called "druggable genome"[9] or the whole genome.

The most widely used RNAi technology in cultured mammalian cells is the introduction of chemically synthesised siRNAs of 21 nucleotides length with symmetric 3′-overhangs of two nucleotides.[10] Due to repeated observations of severe off-target effects[11–17] there is a trend to introduce chemical modifications to inactivate the passenger (sense) strand and/or the seed region of the guide (antisense) strand.[18–21] An alternative approach involves the production of short hairpin siRNAs (shRNA) that are transcribed from expression cassettes inserted in plasmids or viral vectors, that enables longer-term RNAi treatment or even stable integration of such expression cassettes in the treated cells.[22,23] Yet another approach mimics the natural mechanism employed by the model organisms *Caenorhabditis elegans* and *Drosophila*, whereby heterogeneous mixes of active RNAi molecules are produced by endoribonuclease digestion of a larger precursor. Such libraries of endoribonuclease-derived siRNAs (esiRNA) are produced *in vitro* from DNA templates transcribing long double stranded RNA that is subsequently digested by Dicer or RNaseIII to a pool of overlapping effector molecules. Each individual siRNA is present in the pool at very low concentration, so that putative off-target effects of individual molecules may be less likely. Further, the generation of such libraries, which is possible within one's own research lab might be a financially interesting opportunity for academic research labs.[24–26]

Elaborating an appropriate screening strategy

Applying chemically synthesized siRNAs or esiRNAs provides a transient silencing effect. Degradation of the target mRNA is usually maximal after, at the latest, 24 h and recovers by 96 h. Consequently, a timely screening window between 48 and 96 h seems to be appropriate. Unfortunately, proteins exhibiting longer half lives might not show a phenotype within the proposed time frame and others might significantly affect cell viability when depleted after only a short period of time. These considerations have to be taken into account when determining the incubation time between siRNA transfer and assay read-out.

siRNAs are a powerful tool for functional gene analysis and are easily applied for large-scale screens. On the other hand, we and many others have learned of the severe risk of receiving false-positive hits due to off-target effects. We have further observed that negative effects on cell viability and/or the induction of apoptosis as a consequence of target protein depletion may frequently interfere with the desired results of the assay. Statistical analysis might identify this unwanted complication. Therefore, it is of extreme importance to define a proper screening and hit verification strategy up-front, particularly as the strategy might have tremendous financial consequences. A set of minimal requirements for RNAi studies have been outlined early on in the introduction of the technology by a Nature Cell Biology Editorial,[27] and further specified subsequently, particularly for large-scale experiments.[28–31]

Currently, we observe two mainstreams, those who run screens applying pools of 3–4 individual non-overlapping siRNAs against the same target in one well, and those who run 2–4 individual siRNAs in separate wells (Fig. 1). Substances are applied typically in triplicates. In pools, the risk of unwanted off-target effects is lowered as the concentration of the individual siRNAs is decreased while the total amount of siRNA molecules targeting the desired mRNA stays identical. The alternative strategy applies the individual siRNAs in separate wells and, to keep the costs neutral, may skip the replicates. Typically in high-content screening a number of imaging frames, well distributed across the plain of the well are taken, which represent replicates of an identically treated cell population. The variation between replicates is usually not higher than within the frames of one well. At least 200 cells should be covered, but depending on the assay this number might need to be significantly exceeded. The application of individual siRNAs exhibits the advantage that three (or four) individual sequences should induce the identical phenotype. In our labs we are applying 3 siRNAs against each

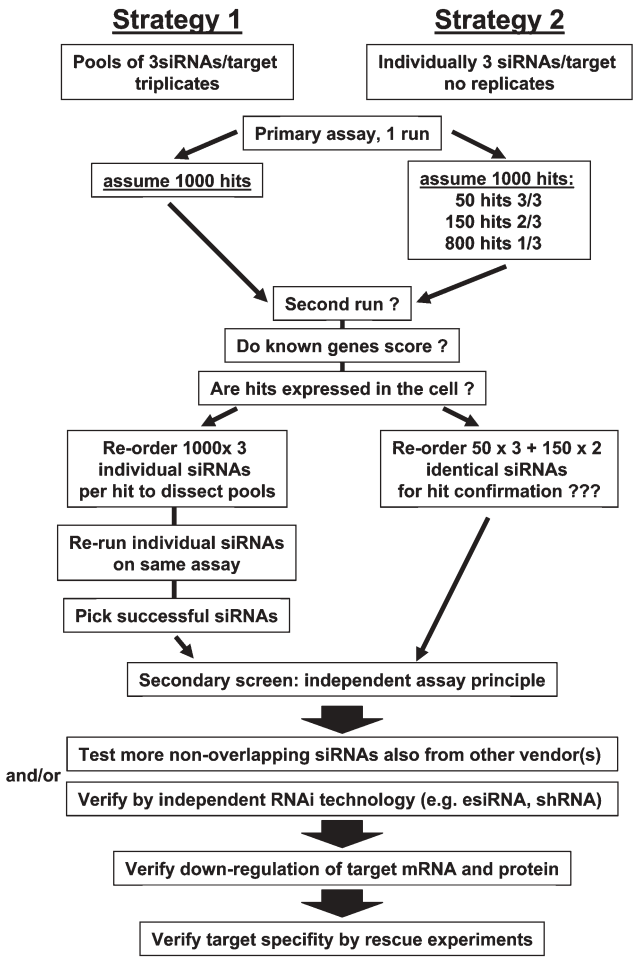

Fig. 1 Considerations for siRNA screening and hit verification strategies.

target. Three out of three (3/3) resulting in the same phenotype is rarely observed because the algorithms generate siRNAs with a probability of 80–90% reducing >70% target mRNA knock-down. This unavoidably means that a certain number of siRNAs is ineffective or exhibits lower potency. Furthermore, a number of phenotypes only occur if the target is depleted to a far greater degree, which might only be achieved by a smaller percentage of the designed siRNAs. More frequently 2/3 identical phenotypes will be observed. Unfortunately, a high number of 1/3-hits will be generated. This ratio raises the suspicion that the one molecule might exhibit an off-target effect.

A second independent run of the whole screen is probably advisable to confirm the first data set and reduce the number of false-positives. Comparing the hits with gene expression array data for the cell line employed in the screen might be helpful to filter the high primary hit number. If the screen was performed applying pools, definitively the next step is to dissect the pools and determine the active siRNA molecules - which might be a cash-intensive step! If individual siRNAs were applied, only the active siRNAs have to be re-ordered; and the 1/3 hits might be rejected. Re-ordering should be done giving sequences rather than siRNA ID numbers so that potential sequence errors might be detected as the siRNAs themselves can hardly be sequenced. It is a reasonable question to ask whether re-ordering of the identical molecules is at all meaningful rather than immediately proceeding to ordering additional non-overlapping siRNAs possibly from a different vendor. Alternatively or additionally, another RNA interference technology might be applied, for example endoribonuclease-derived siRNAs or short hairpin siRNAs. For the derived solid hits and molecules, target mRNA down-regulation by RT-PCR and protein depletion by ELISA, Western or immunostaining in cells should be demonstrated, if an antibody is available against a given target. Finally, rescue experiments in cells should demonstrate that the phenotype is really induced by depletion of the targeted protein. These rescue experiments might be achieved by co-transfected cDNA of a different organism, or by introducing siRNA-resistant mutants or a very elegant method involves making a stable cell line that has in addition to the endogenous gene of interest, a bacterial artificial chromosome (BAC) inserted that expresses the corresponding mouse gene from its upstream 100–150 kb endogenous promoter sequence.[32,33] This cell model allows selective down-regulation of the endogenous human, the exogenous mouse or both genes by appropriately designed siRNA molecules, and should manifest a clear phenotype only upon depletion of both genes.

How to deliver siRNA molecules efficiently and at appropriate throughput?

A broad range of delivery agents particularly to transfect siRNAs has been developed and commercialised. Most systems are lipid-based, but alternatives such as polymer- or peptide-based reagents have been introduced. A permanently updated list with links to the providers resides at www.biocompare.-com. Most reagents work excellently for easy to transfect cell lines such as HeLa, COS-7, MCF-7, A549, HEK 293 or DU145, but the majority of interesting cell models is more difficult to treat and a range of reagents have to be tested to gain success. For higher throughput, automated transfection methods have been developed in 96-well microtiter plate format that allow large-scale siRNA screening. Most published protocols are established at a final siRNA concentration of 100 nM on the cells which runs into severe risk of inducing off-target effects.[34,35] Reverse liquid transfection (place the transfection complex in the well first, then add the cells on top) has been demonstrated to achieve higher transfection efficiency at far lower siRNA concentrations than traditional forward transfection (cells seeded the day before, and transfection complexes are administered to adherent cell layers).[36] In

our lab protocols have been established applying 96 or 384 steel or Teflon-coated needles and performing stringent washing and disinfection protocols to avoid plastic tips completely and that results in substantial cost savings during large scale screening campaigns. Further, final siRNA concentrations on the cells could be brought down to the low or even sub-nanomolar range. Five 384-well library plates are run in parallel per 50 min achieving a throughput of the whole genome in 29 to 49 h, the latter with a library covering 23 000 genes and applying individually each 4 siRNAs per gene (unpublished data). Alternatively, impressive progress has been achieved with reverse solid phase high-throughput transfection using slides of spotted transfection complexes in an arrayed format.[37–39]

Viral vector systems[8,22,23] offer the benefit of high infection rates that may be particularly interesting for more difficult to transfect cells such as primary cells that will gain more and more importance in future screening using physiologically more relevant systems. Vector-based libraries comprise the beauty of an unlimited resource provided one has the high-throughput vector production feasibility that requires for viral vector production a biosafety level 2 certification.[8] On the other hand, the stability of the viral stocks may be limiting. Repeated freeze/thaw cycles are not well tolerated by viruses, and therefore a purchased ready-to-use library might have a very limited life-span.

Very recently, electroporation has been brought to a 96-well format and holds promise for primary and suspension cell screening at a higher scale. Application of milder conditions is claimed to allow successful siRNA traversation across the membrane into the cytosol than compared to plasmid delivery that require expression in the nucleus. However, the mechanism of electroporation is not well understood, therefore developing protocols empirically is requested by adjusting pulse parameters such as voltage (amplitude), pulse duration, number of pulses and duration of inter-pulse intervals. Cells need to be suspended in low-conductivity buffers that unfortunately affect cell viability negatively and allow handling for less than an hour without inducing significant cell damage. Interestingly, a recent publication demonstrated significantly less off-target effects induced by electroporating siRNAs compared to lipofection measuring gene expression by microarrays.[40] The siPORTer™-96 Electroporation Chamber (Ambion, Austin, TX, USA) offers a very simple solution for 96-well electroporation. Two re-usable electrode plates with 96 cavities each form shells like a mussel and harbour the drops of 40–50 μL volume consisting of cells and siRNAs in an electroporation buffer.[36] After use the electrode plates can be moved out and cleaned for re-use simply by rinsing with 70% ethanol. Conditions are published that consume 1 μg (1000 nM on the cells!) or 1.5 μg siRNA per reaction. With this system it is very difficult to optimise transfection conditions as there is only one setting per plate applicable. Further concerns are the risk of cross-contamination and putatively uneven distribution of the conditions across the plate resulting in higher variation than other systems. The PA-4000/PA-96WS PulseAgile 96-Well Electroporation System (CytoPulse Sciences, Glen Burnie, MD, USA) offers electroporation in low cost polypropylene microplates by arrayed re-usable coaxial electrodes lowered into the samples for pulsing. The system allows addressing any row, any column, or all wells concurrently. Currently, three protocols are published. Electroporation is performed at 100 or 800 nM siRNA concentration in 100 or 200 μL at 1×10^5 to 1×10^6 RAW264.7, PC-12 or Jurkat cells [www.cytopulse.com/lab-research/protol.shtml[41]]. The electrode array can easily cleaned and re-used afterwards. To the next generation of instruments belong the BTX HT 96 Well Electroporation System (BTX Molecular Delivery Systems, San Diego, CA, USA)[42] and the Nucleofector 96-well Shuttle System (amaxa, Cologne/Köln, Germany). Both systems provide disposable multi-well plates with integrated electrodes. The reaction volume and therefore the required siRNA and cell amounts are significantly lowered, which is particularly for primary cells a critical factor. BTX plates cover maximum volumes of 125 μL or 250 μL depending on gap size, while amaxa shows data to reduce the reaction volume to a minimum of 20 μL at variations of 1–3% CV. Each well can be addressed individually and allows full flexibility to optimise assay conditions while easy accessibility will allow automation for higher throughput. Both systems hold strong promises although the concentration of siRNAs and consumption of cells needs further optimisation while the costs for consumables will need to decline significantly in order for the technology to gain broader application.

Which is the right image acquisition instrument?

The variety of available image acquisition instruments is broad, and has to carefully be assessed against the specific needs of the screening operation and the available budget. Laser scanning fluorimeters, such as the Acumen Explorer or the Blueshift IsoCyte, offer high-speed at low resolution and simple reader-like handling. Laser scanning excitation at various wavelengths is combined with photomultiplier tube detection, that allows detection of up to 4 colours simultaneously and a large field of view (20×20 mm^2) means the analysis is performed per field not per well. This brings the image acquisition time of a whole plate down to a few minutes regardless of the plate format (96, 384, 1536 wells per plate). The drawback is that the resolution is not very high but is nevertheless sufficient to allow analysis of fluorescent intensity, size, location or morphological changes at the cellular level, which may be sufficient for many primary screens. Popular assays such as proliferation, mitotic index, cytotoxicity, apoptosis, reporter gene expression, translocation between cytosol and nucleus or neurite outgrowth can be performed well applying these instruments.

In developing sophisticated high-speed microscopy systems the major achievements include robustness of autofocus, rapid local sub-micron positioning even when the same segment is re-visited later, stable light sources, higher resolution by water immersion devices, and using objectives at higher NAs that catch more light and allow greater depth of field. Besides fluorescent microscopy, some systems support bright field/phase

contrast microscopy for non-stained cells. Further equipment components have been added such as injection devices and climatization chambers, or components to allow the integration into a core IT infrastructure or interfacing with lab automation, which has been successfully proven. Plastic bottom microtiter plates have been produced that reach imaging quality similar to the much more expensive glass bottom plates that had to be used previously.

Probably the most widely used automated epifluorescent microscopy system currently is the ArrayScan of the HCS microscopy market leader, Cellomics. The latest ArrayScan VTI model offers rapid automated autofocusing, one 12 bit CCD camera, a broad excitation spectrum provided by a mercury-xenon lamp, automated filter wheels for excitation and emission, and four lenses at $5\times$, $10\times$, $20\times$ and $40\times$ magnification. As multi-coloured images have to be acquired in sequence (rather than simultaneously) at the individual wavelengths the throughput is limited and typically reaches about 10 000 images per day. A *pseudo*-confocality may be optionally provided by the Zeiss ApoTome. Further, the system is fitted with a data base, and is ready to be upgraded with a selection of image analysis packages referred to as BioApplications. Along with ready-to-use biological kits this system offers a turnkey solution. Competitors have followed by providing similar complete systems such as the IN Cell Analyzer 1000 of GE Healthcare.

More recently, a number of confocal systems has been introduced such as the IN Cell Analyzer 3000 (GE Healthcare, formerly Amersham BioSciences), the Opera (Evotec Technologies), the BD Pathway BD Pathway™ 435 or 855 Bioimager (Becton Dickinson, formerly Atto Bioscience) and the ImageXpress Ultra (Molecular Devices) [ref. 43–45, Table 1]. Confocality provides vertical spatial resolution meaning the cells are sliced visually akin to how a salami is sliced at the butcher. Confocality also reduces background fluorescence (*e.g.* free fluorophores or autofluorescent compounds) contributed to by out-of-focus fluorescence in the supernatant or due to interference from adjacent object features above or below the focal plane. Under certain assay conditions confocality may provide further increases in sensitivity by specific enrichment for a desired phenotype such as has been shown for apoptotic cells that detach from the growth surface and round up (Fig. 2).

We have been using the OPERA system (Evotec Technologies, Hamburg, Germany) for a number of large-scale screening projects. The system offers excellent image quality, for example due to the true confocality provided by a Nipkow spinning disk or the water-immersion system employed to increase sub-cellular resolution. The latest version of the OPERA QEHS is illuminated by four solid state lasers at 405, 488, 532 and 635 nm as well as a Xenon high pressure lamp to provide a non-confocal UV mode. Three cooled CCD cameras allow simultaneous acquisition in three colours. Additionally, one further camera is fitted for the UV mode. An automated objective changer allows image acquisition at $10\times-60\times$ (air objectives) or $20\times-60\times$ (water-immersion objectives). The system can be upgraded to integrate a climatization chamber providing a controlled temperature, CO_2 and humidity environment. Further, a dispensing unit exhibiting a single needle allows injection of reagent volumes as little as 0.5 μL for example to analyse Ca^{2+}-release immediately after external stimulation.

Furthermore, all of these confocal imaging systems provide the user with integrated image analysis tools and sophisticated software algorithms to allow the analysis of cellular fluorescence and the determination of morphological characteristics. Some also offer interfaces to external image analysis software.

Image analysis, a challenging and quickly emerging field

An essential factor in the success of high-content screening projects is the existence of algorithms and software that can reliably and automatically extract information from the masses of captured images. Typically, nuclei are identified and masked first. Then, areas around the nuclei are determined or the cell boundaries are searched to mask the cell shape. Dyes that stain not only chromosomal DNA in the nucleus but also mRNA in the cytosol help to identify the cell's shape. Nuclei may be counted along with extraction of additional parameters such as shape, size, substructures like spots, or intensity. Subsequently, the masks are laid over the image(s) of the other channel(s) and signals within the masks are measured (Fig. 3). As mentioned above, most of the more advanced automated microscopes are delivered with proprietary image analysis solutions for a broad range of biological events. For popular assays at the sub-cellular level such as cell cycle analysis (mitotic index), cytotoxicity, apoptosis, micronuclei detection, receptor internalisation, protein translocation (membrane to

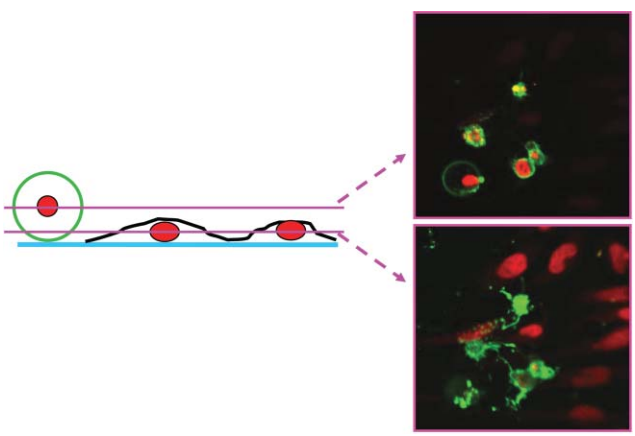

Fig. 2 Confocal microscopy not only increases image quality but as well may increase sensitivity towards the desired phenotype. For the Draq5/AnnexinV cell proliferation/apoptosis screen, performing the imaging at a confocal plane above the substrate allows specific detection of the desired phenotype while leaving the healthy unaffected adherent cells below the imaging plane. Therefore, the screening window could be largely increased.

Table 1 Comparison of confocal high-content screening systems

	OPERA QEHS	IN Cell Analyser 3000	Pathway Bioimager	Image Express Ultra
Provider	Evotec Technologies	GE Healthcare	BD	Molecular Devices
Light sources	405, 488, 535, 635 nm, solid state lasers	364, 488, 647 nm, ion lasers	Dual HBO 100W mercury arc bulb (full spectrum 340–740 nm)	Up to 4 lasers 405, 488, 532, 635 nm, solid state lasers
UV	Yes, non-confocal, +separate camera	Yes, 364 nm laser	Yes	Yes
Laser autofocus	Yes	Yes	Yes	Option
Image-based autofocus	—	—	Yes	Yes
Confocal optics	Nipkow spinning disc	Line scanning	Nipkow spinning disc, can be moved in and out	Point scanning
Magnifications	4×–60×	40× fixed, large view field (750 × 750 μm)	Olympus 4×, 10×, 20×, 40×, 60×	Nikon 4×–100×, 4 lenses
Water-immersion	Yes, 20×–60×	No	No	oil-immersion, 40×–100×
Detection systems	3 (4) CCD cameras	3 CCD	1 CCD	4 photomultipliers, sequential or parallel
Liquid handling	Yes	Yes	Yes	No
Incubation chamber	CO2, temperature, humidity	CO2, temperature, humidity	CO2, temperature	No
Bright field option	No	No	Yes	No
Binocular eyepieces	No	No	Yes	No
Plate formats	6–1536 well microtiter plates, microscopy slides	96 & 384 well microtiter plates	96/384 well microtiter plates, microscopy slides	6–1536 well microtiter plates, microscopy slides

cytosol, cytosol to nucleus, and *vice versa*), co-localisation, cytoskeletal arrangements, or morphological analysis at the cellular level such as neurite outgrowth, cell spreading, cell motility, colony formation, or tube formation, ready-to-use scripts are available and need only some fine-adjustment for the particular cell line and/or conditions of the assay. Besides the packs provided by the microscope suppliers, a number of commercial products are available that might be used alternatively. We have successfully employed the Cellenger software (Definiens, Munich/München, Germany) for off-line analysis that offers an object-oriented approach to analyse complex phenotypes more precisely but at lower speed.[46,47]

An academic consortium started by Anne E. Carpenter and Thouis R. Jones in the laboratories of David M. Sabatini and Polina Golland at the Whitehead Institute for Biomedical Research and MIT's CSAIL has been developing an open-source cell image analysis software called CellProfiler™ that is designed for biologists without training in computer vision or programming to quantitatively measure cellular phenotypes in thousands of images automatically (www.cellprofiler.org). New modules can be added using Matlab. Alternatively, the Open Microscopy Environment (OME) launched by the University of Dundee[48–50] is continuously developed further towards the needs of image management and analysis in support of high-content screening as well as traditional microscopy.

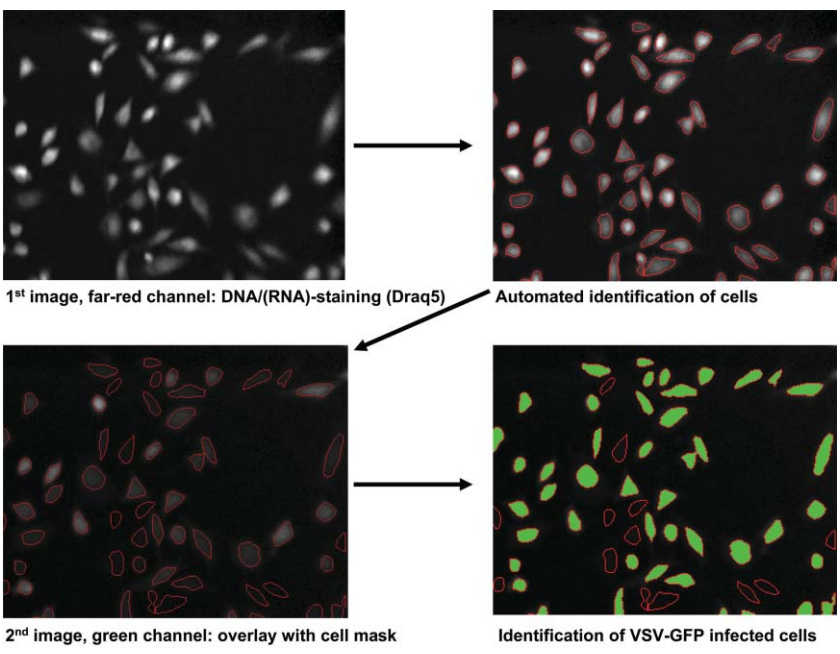

Fig. 3 Automated image analysis: counting cells that are infected by a GFP-recombinant virus. Images were acquired by the OPERA TEHS and analysed on-the-fly by Acapella™[45] (Evotec Technologies, Hamburg, Germany).

What to do with all the data?

As mentioned earlier, high-content screening can easily generate more than one Terabyte in primary images and metadata per run, that have to be stored and organised, which means an appropriate laboratory information management system (LIMS) has to be established.[51] The LIMS must be able to collect, collate and integrate the data stream to allow at least searching and rapid evaluation of the data. After image acquisition and data transfer, image analysis will be run to extract the metadata. Further evaluation includes testing for process errors. Heat maps along with pattern recognition

algorithms help to identify artefacts such as edge-effects, uneven pipetting, or simply to exclude images that are not in focus. All plates should be checked so that the selected positive and negative controls exhibit values in a pre-defined range. Further, data may be normalised against controls before further statistical analysis is run to identify putative hits. Known proteins of the pathway being screened should score, and are a good internal control for the accuracy of the assay and workflow. Hits have to be verified by going back to the original images. Further, results have to be compared between independent runs. After this, an appropriate hit verification strategy has to be applied as discussed above. Target gene expression should be confirmed, for example, by running a microarray analysis of gene expression for the given cell line. Finally, data will be compared to other internal and external data sources. Cluster analysis will assist in identifying networks and correlations.

Examples of large-scale siRNA screens

Meanwhile, the literature is rich with large-scale RNAi screens in cultured mammalian cells, and the list is growing rapidly. Among the first published screens different RNAi technologies have been proven such as retrovirally expressed shRNAs,[52] endoribonuclease-prepared (e)siRNAs[53] or chemically synthesized siRNAs.[54] To review all screens would be the subject of another article. An important application of large-scale siRNA screening is to sensitize cells to current chemotherapy[55,56] and will help the pharmaceutical industry and the clinicians to identify additional targets for new drug discovery programmes and to find new combination treatments that will hopefully accelerate response rates for example in fighting cancer.

Outlook

Although tremendous progress has been achieved in HCS, there are major challenges remaining to be solved. More complex and physiologically more relevant assays will be developed in primary cells, multi-cellular systems, organ cultures or whole model organisms. Further, there is increased interest to visualise and screen complex dynamic processes in living cells by recording and analysing movies.[57] Therefore, image analysis has to be improved to extract reliably and in a well quantifiable manner (far) more parameters from assays of increasing complexity. Object tracking in movies is a current hot-spot of research in automated image analysis.

A further topic of increasing importance is data mining. Relationships have to be generated between one's own extracted rich data and external knowledge by bioinformatics tools, and may also require battling with conflicting data from literature of unknown or difficult to assess quality, and respecting the intrinsic variability of cell-based biological assays. Functional genomic (RNAi) screening will be extended by chemical screening to derive probes to block individual protein functions rather than depleting the whole protein as achieved by gene silencing. Merging of functional and chemical genomics data will derive patterns that will dramatically broaden our understanding of complex biological processes.

Acknowledgements

Marino Zerial and Ivan Baines, both directors at the MPI-CBG, provided the vision and concepts at a very early time point. Thanks to a highly motivated, skilled and dedicated multi-disciplinary team a unique infrastructure has been built-up and established that allows implementation of world-class high content RNAi assay development and screening projects together with our scientific and technology partners. Last not least, funding by the BMBF/InnoRegio/BioMeT grant (Förderkennzeichen 03I4035A) and the Max Planck Society's inter-institutional initiatives "RNA Interference" and "Chemical Genomics Centre" allowed setting-up and operation of the TDS as well as implementation of a significant number of its projects. Very special thanks belong to Dr Marjo Simonen, Lab Head at Novartis, Basel, who contributed an outstanding overview on current automated imaging systems to the RIGHT-Symposium in Dresden in September 2006 that provided the basis for the paragraph in this review.

References

1 K. A. Giuliano, R. L. DeBiasio, R. T. Dunlay, A. Gough, J. M. Volosky, J. Zock, G. N. Pavlakis and D. L. Taylor, High Content Screening: A new approach to easing key bottlenecks in the drug discovery process, *J. Biomol. Screen.*, 1997, **2**, 249–259.

2 B. R. Stockwell, S. J. Haggarty and S. L. Schreiber, High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications, *Chem. Biol.*, 1999, **6**(2), 71–83.

3 Y. Tomari and P. D. Zamore, Perspective: machines for RNAi, *Genes Dev.*, 2005, **19**(5), 517–29.

4 D. M. Dykxhoorn, C. D. Novina and P. A. Sharp, Killing the messenger: short RNAs that silence gene expression, *Nat. Rev. Mol. Cell Biol.*, 2003, **4**(6), 457–67.

5 Y. Pei and T. Tuschl, On the art of identifying effective and specific siRNAs, *Nat. Methods*, 2006, **3**(9), 670–6.

6 K. Ui-Tei, Y. Naito and K. Saigo, Guidelines for the selection of effective short-interfering RNA sequences for functional genomics, *Methods Mol. Biol.*, 2006, **361**, 201–16.

7 W. Gong, Y. Ren, Q. Xu, Y. Wang, D. Lin, H. Zhou and T. Li, Integrated siRNA design based on surveying of features associated with high RNAi effectiveness, *BMC Bioinformatics*, 2006, **7**, 516.

8 D. R. Rines, B. Tu, L. Miraglia, G. L. Welch, J. Zhang, M. V. Hull, A. P. Orth and S. K. Chanda, High-content screening of functional genomic libraries, *Methods Enzymol.*, 2006, **414**, 530–65.

9 A. L. Hopkins and C. R. Groom, The druggable genome, *Nat. Rev. Drug Discovery*, 2002, **1**(9), 727–30.

10 S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschl, Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells, *Nature*, 2001, **411**(246836), 494–8.

11 D. Semizarov, L. Frost, A. Sarthy, P. Kroeger, D. N. Halbert and S. W. Fesik, Specificity of short interfering RNA determined through gene expression signatures, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**(11), 6347–52.

12 A. L. Jackson, S. R. Bartz, J. Schelter, S. V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet and P. S. Linsley, Expression profiling reveals off-target gene regulation by RNAi, *Nat. Biotechnol.*, 2003, **21**(6), 635–7.

13 P. C. Scacheri, O. Rozenblatt-Rosen, N. J. Caplen, T. G. Wolfsberg, L. Umayam, J. C. Lee, C. M. Hughes, K. S. Shanmugam, A. Bhattacharjee, M. Meyerson and F. S. Collins, Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**(7), 1892–7.

14 L. P. Lim, N. C. Lau, P. Garrett-Engele, A. Grimson, J. M. Schelter, J. Castle, D. P. Bartel, P. S. Linsley and

J. M. Johnson, Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs, *Nature*, 2005, **433**(7027), 769–73.
15 X. Lin, X. Ruan, M. G. Anderson, J. A. McDowell, P. E. Kroeger, S. W. Fesik and Y. Shen, siRNA-mediated off-target gene silencing triggered by a 7 nt complementation, *Nucleic Acids Res.*, 2005, **33**(14), 4527–35.
16 A. Birmingham, E. M. Anderson, A. Reynolds, D. Ilsley-Tyree, D. Leake, Y. Fedorov, S. Baskerville, E. Maksimova, K. Robinson, J. Karpilow, W. S. Marshall and A. Khvorova, 3′ UTR seed matches, but not overall identity, are associated with RNAi off-targets, *Nat. Methods*, 2006, **3**(3), 199–204.
17 A. L. Jackson, J. Burchard, J. Schelter, B. N. Chau, M. Cleary, L. Lim and P. S. Linsley, Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity, *RNA*, 2006, **12**(7), 1179–87.
18 F. Czauderna, M. Fechtner, S. Dames, H. Aygun, A. Klippel, G. J. Pronk, K. Giese and J. Kaufmann, Structural variations and stabilising modifications of synthetic siRNAs in mammalian cells, *Nucleic Acids Res.*, 2003, **31**(11), 2705–16.
19 T. P. Prakash, C. R. Allerson, P. Dande, T. A. Vickers, N. Sioufi, R. Jarres, B. F. Baker, E. E. Swayze, R. H. Griffey and B. Bhat, Positional effect of chemical modifications on short interference RNA activity in mammalian cells, *J. Med. Chem.*, 2005, **48**(13), 4247–53.
20 B. A. Kraynack and B. F. Baker, Small interfering RNAs containing full 2′-O-methylribonucleotide-modified sense strands display Argonaute2/eIF2C2-dependent activity, *RNA*, 2006, **12**(1), 163–76.
21 A. L. Jackson, J. Burchard, D. Leake, A. Reynolds, J. Schelter, J. Guo, J. M. Johnson, L. Lim, J. Karpilow, K. Nichols, W. Marshall, A. Khvorova and P. S. Linsley, Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing, *RNA*, 2006, **12**(7), 1197–205.
22 J. M. Silva, M. Z. Li, K. Chang, W. Ge, M. C. Golding, R. J. Rickles, D. Siolas, G. Hu, P. J. Paddison, M. R. Schlabach, N. Sheth, J. Bradshaw, J. Burchard, A. Kulkarni, G. Cavet, R. Sachidanandam, W. R. McCombie, M. A. Cleary, S. J. Elledge and G. J. Hannon, Second-generation shRNA libraries covering the mouse and human genomes, *Nat. Genet.*, 2005, **37**(11), 1281–8.
23 D. E. Root, N. Hacohen, W. C. Hahn, E. S. Lander and D. M. Sabatini, Genome-scale loss-of-function screening with a lentiviral RNAi library, *Nat. Methods*, 2006, **3**(9), 715–9.
24 D. Yang, F. Buchholz, Z. Huang, A. Goga, C. Y. Chen, F. M. Brodsky and J. M. Bishop, Short RNA duplexes produced by hydrolysis with Escherichia coli RNase III mediate effective RNA interference in mammalian cells, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**(15), 9942–7.
25 R. Kittler, A. K. Heninger, K. Franke, B. Habermann and F. Buchholz, Production of endoribonuclease-prepared short interfering RNAs for gene silencing in mammalian cells, *Nat. Methods*, 2005, **2**(10), 779–84.
26 F. Buchholz, R. Kittler, M. Slabicki and M. Theis, Enzymatically prepared RNAi libraries, *Nat. Methods*, 2006, **3**(9), 696–700.
27 Editorial: Whither RNAi?, *Nat. Cell. Biol.*, 2003, **56**, pp. 489–90.
28 E. Krausz, A. Grabner, A. Kroenke, C. Sachse and C. J. Echeverri, Optimising high throughput RNAi-based assays using transient transfection of synthetic siRNAs in cultured mammalian cells, in *RNA Interference (RNAi)- Nuts & Bolts of RNAi Technology*, ed. D.R. Engelke, DNA Press LLC, 2003, pp. 131–168.
29 C. Sachse, E. Krausz, A. Kronke, M. Hannus, A. Walsh, A. Grabner, D. Ovcharenko, D. Dorris, C. Trudel, B. Sonnichsen and C. J. Echeverri, High throughput RNA interference strategies for target discovery and validation using synthetic short interfering RNAs: functional genomics investigations of biological pathways, *Methods Enzymol.*, 2005, **392**, 242–77.
30 C. J. Echeverri and N. Perrimon, High-throughput RNAi screening in cultured cells: a user's guide, *Nat. Rev. Genet.*, 2006, **7**(5), 373–84.
31 C. J. Echeverri, P. A. Beachy, B. Baum, M. Boutros, F. Buchholz, S. K. Chanda, J. Downward, J. Ellenberg, A. G. Fraser, N. Hacohen, W. C. Hahn, A. L. Jackson, A. Kiger, P. S. Linsley, L. Lum, Y. Ma, B. Mathey-Prevot, D. E. Root, D. M. Sabatini, J. Taipale, N. Perrimon and R. Bernards, Minimizing the risk of reporting false positives in large-scale RNAi screens, *Nat. Methods*, 2006, **3**(10), 777–9.
32 R. Kittler, L. Pelletier, C. Ma, I. Poser, S. Fischer, A. A. Hyman and F. Buchholz, RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**(7), 2396–401.
33 I. Poser and F. Buchholz, A new approach to specify RNAi experiments, *Eur. Pharm. Rev.*, 2006, **3**, 12–17.
34 N. Machuy, B. Thiede, K Rajalingam, C. Dimmler, O. Thieck, T. F. Meyer and T. Rudel, A global approach combining proteome analysis and phenotypic screening with RNA interference yields novel apoptosis regulators, *Mol. Cell. Proteomics*, 2005, **4**(1), 44–55.
35 K. G. Moore, W. Speckmann and R. P. Herzig, The use of siRNA to validate immunofluorescence studies, *Methods Mol. Biol.*, 2007, **356**, 245–51.
36 D. Ovcharenko, R. Jarvis, S. Hunicke-Smith, K. Kelnar and D. Brown, High-throughput RNAi screening in vitro: From cell lines to primary cells, *RNA*, 2005, **11**, 985–993.
37 H. Erfle, J. C. Simpson, P. I. Bastiaens and R. Pepperkok, siRNA cell arrays for high-content screening microscopy, *Biotechniques*, 2004, **37**(3), 454–462.
38 D. B. Wheeler, A. E. Carpenter and D. M. Sabatini, Cell microarrays and RNA interference chip away at gene function, *Nat. Genet.*, 2005, **37 Suppl**, S25–30.
39 B. Neumann, M. Held, U. Liebel, H. Erfle, P. Rogers, R. Pepperkok and J. Ellenberg, High-throughput RNAi screening by time-lapse imaging of live human cells, *Nat. Methods*, 2006, **3**(5), 385–90.
40 Y. Fedorov, A. King, E. Anderson, J. Karpilow, D. Ilsley, W. Marshall and A. Khvorova, Different delivery methods-different expression profiles, *Nat. Methods*, 2005, **2**(4), 241.
41 R. Jarvis and A. D. King, Optimizing Chemical Transfection and Electroporation of siRNAs, *Ambion TechNotes*, 2003, **10**(5), 12–15.
42 W. J. Buchser, J. R. Pardinas, Y. Shi, J. L. Bixby and V. P. Lemmon, 96-well electroporation method for transfection of mammalian central neurons, *Biotechniques*, 2006, **41**(5), 619–24.
43 L. Zemanová, A. Schenk, M. J. Valler, G. U. Nienhaus and R. Heilker, Confocal optics microscopy for biochemical and cellular high-throughput screening, *Drug Discovery Today*, 2003, **8**, 1085–1093.
44 A. H. Gough and P. A. Johnston, Requirements, features, and performance of high content screening platforms, *Methods Mol. Biol.*, 2007, **356**, 41–61.
45 R. J. Garippa, A. F. Hoffman, G. Gradl and A. Kirsch, High-throughput confocal microscopy for beta-arrestin-green fluorescent protein translocation G protein-coupled receptor assays using the Evotec Opera, *Methods Enzymol.*, 2006, **414**, 99–120.
46 M. Baatz, N. Arini, A. Schape, G. Binnig and B. Linssen, Object-oriented image analysis for high content screening: detailed quantification of cells and sub cellular structures with the Cellenger software, *Cytometry A.*, 2006, **69**(7), 652–8.
47 P. Biberthaler, M. Athelogou, S. Langer, B. Luchting, R. Leiderer and K. Messmer, Evaluation of murine liver transmission electron micrographs by an innovative object-based quantitative image analysis system (Cellenger), *Eur. J. Med. Res.*, 2003 Jul 31, **8**(7), 275–82.
48 D. A. Schiffmann, D. Dikovskaya, P. L. Appleton, I. P. Newton, D. A. Creager, C. Allan, I. S. Nathke and I. G. Goldberg, Open microscopy environment and findspots: integrating image informatics with quantitative multidimensional image analysis, *BioTechniques*, 2006, **41**(2), 199–208.
49 I. G. Goldberg, C. Allan, J. M. Burel, D. Creager, A. Falconi, H. Hochheiser, J. Johnston, J. Mellen, P. K. Sorger and J. R. Swedlow, The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging, *Genome Biol.*, 2005, **6**(5), R47.
50 J. R. Swedlow, I. Goldberg, E. Brauner and P. K. Sorger, Informatics and quantitative analysis in biological imaging, *Science*, 2003, **300**(5616), 100–2.

51 K. Kozak, M. Kozak, J. Wagner, H. Grabner, K. Korn, E. Fava, M. Biesold, C. Möbius, A. Lohmann and E. Krausz, LIMS TDS: A Platform for Comprehensive Management and Analysis of Screening Data., *RNAi.net Newsletter*, 2006, **6**(8).
52 K. Berns, E. M. Hijmans, J. Mullenders, T. R. Brummelkamp, A. Velds, M. Heimerikx, R. M. Kerkhoven, M. Madiredjo, W. Nijkamp, B. Weigelt, R. Agami, W. Ge, G. Cavet, P. S. Linsley, R. L. Beijersbergen and R. Bernards, A large-scale RNAi screen in human cells identifies new components of the p53 pathway, *Nature*, 2004, **428**(6981), 431–7.
53 R. Kittler, G. Putz, L. Pelletier, I. Poser, A. K. Heninger, D. Drechsel, S. Fischer, I. Konstantinova, B. Habermann, H. Grabner, M. L. Yaspo, H. Himmelbauer, B. Korn, K. Neugebauer, M. T. Pisabarro and F. Buchholz, An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division, *Nature*, 2004, **432**(7020), 1036–40.
54 L. Pelkmans, E. Fava, H. Grabner, M. Hannus, B. Habermann, E. Krausz and M. Zerail, Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis, *Nature*, 2005, **436**(7047), 78–86.
55 J. P. MacKeigan, L. O. Murphy and J. Blenis, Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance, *Nat. Cell Biol.*, 2005, **7**(6), 591–600.
56 S. R. Bartz, Z. Zhang, J. Burchard, M. Imakura, M. Martin, A. Palmieri, R. Needham, J. Guo, M. Gordon, N. Chung, P. Warrener, A. L. Jackson, M. Carleton, M. Oatley, L. Locco, F. Santini, T. Smith, P. Kunapuli, M. Ferrer, B. Strulovici, S. H. Friend and P. S. Linsley, siRNA Screens Reveal Enhanced Cisplatin Cytotoxicity in Tumor Cells Having Both BRCA Network and TP53 Disruptions, *Mol. Cell. Biol.*, 2006, **26**(24), 9377–86.
57 B. Neumann, M. Held, U. Liebel, H. Erfle, P. Rogers, R. Pepperkok and J. Ellenberg, High-throughput RNAi screening by time-lapse imaging of live human cells, *Nat. Methods*, 2006, **3**(5), 385–90.

